Asian Journal of Information Technology 13 (10): 633-638, 2014
ISSN: 1682-3915

Combined Global-Local Specialized Feature Descriptor for Content Based Image Retrieval under Noisy Query

[1]S. Singaravelan and [2]D. Murugan
[1]PSR Engineering College, Sivakasi, India
[2]Manonmaniam Sundaranar University, Tirunelveli, India

Abstract: The lot of research on Content Based Image Retrieval (CBIR) considering the different image/visual features like color, shape, texture and semantic methods has been done earlier. In the real world environment, the noises that may embed into an image document will affect the CBIR Environments algorithms. Tough different filtering algorithms are available for noise reduction, applying a Filtering algorithm that is sensitive to one type of noise to an image which has been degraded by another type of noise lead to unfavorable results. This condition stresses the importance of designing a efficient CBIR algorithm that retains precision rates even under noisy conditions. In this research, numerous experiments have been conducted to analyze the robustness of the proposed Combined Global-Local Specialized Features Descriptor (CGLSFD). This proposed methods include two stages. First apply wavelet to decompose the query image to extracted the energy, standard deviation and mean values in all bands. Second apply Micro Structure Descriptor (MSD) to extract image edge orientation features with color, texture and shape and color layout information. This proposed method extensively tested on corel data tests. This proposed CGLSFD algorithm has high indexing and low dimensionality also along with other Existing Conventional algorithms under different types of noises such as Gaussian noise, salt and pepper noise and quantization noises. This proposed CGLSFD algorithms results compare with other Existing Conventional algorithms than Gabor features and MSD in image retrieval.

Key words: CBIR, multiwavelet transform, micro structure descriptor, noise reduction, CGLSFD

INTRODUCTION

Content Based Image Retrieval (CBIR) is any technology that in principle helps to organize digital image archives by their visual content. By this definition, anything ranging from an image similarity function to a robust image annotation engine falls under the purview of CBIR. The most common form of CBIR is an image search based on visual. The increasing amount of digitally produced images requires new methods to archive and access this data. Conventional databases allow for textual searches on Meta data only. Content Based Image Retrieval (CBIR) is a technique which uses visual contents, normally called as features, to search images from large scale image databases according to users' requests in the form of a query image. Apart from the usual features like color and texture, a new Feature Extraction algorithm called CGLSFD introduced. CGLSFD convey essential information to a picture and therefore can be applied to image retrieval. The CGLSFD captures the spatial distribution of edges, color layout, shape and texture and multi-wavelet low and high frequency bands.

This model expects the input as Query by Example (QBE) and any combination of features can be selected for retrieval. The focus is to build a universal CBIR System using low level features (Zhang and Lu, 2003; Alnihoud, 2012).

LITERATURE REVIEW

CBIR using wavelet transform: The wavelet transform goes further than the short time fourier transform. It's also analyzes the image by multiplying it by a window function and performing an orthogonal expansion, analogously to other linear integral transformation. There are two direction in which they analysis is extended. First direction, the basic functions (called wavelets, meaning a small wave or mother wavelets) are more complicated than sines and cosines. They provide localization in space to a certain degree, not entire localization due to the certainty principal. Second direction, the analysis is performed at multiple scales. Localization in the spatial domain together with the wavelet' localization infrequency yields a sparse representation of many

Corresponding Author: S. Singaravelan, PSR Engineering College, Sivakasi, India

practical images. This sparseness opens the door to successful application in image compression, noise filtering and detecting features in images. Wavelets have been used with enormous success in data compression, feature detection and in image noise suppression, it's possible to erase to zero the contribution of wavelet components that are small and correspond to noise without erasing the important small details in the underlying image (Sonka *et al*., 2008; Do and Vetterli, 2002; Daubechies, 1990; Mallat, 1989).

CBIR using different descriptor: The human visual attention is enhanced through a process of competing interactions among neurons which selects a few elements of attention and suppresses irrelevant materials (Desimone, 1998). There are close relationships between low-level visual features and human visual attention system and hence the research on how to use visual perception mechanism for image retrieval is an important yet challenging problem. In order to extract features via simulating visual processing procedures and effectively inte-grate color, texture, shape features and image color layout information as a whole for image retrieval, in this study, researchers propose a novel feature detector and descriptor, namely Micro-Structures Descriptors (MSD), to describe image features via micro-structures.

The data gathered above have been fed into worksheets to perform statistical analysis on the results in group-projects, project's individual reports and exams. In addition, the results of students' attainment in the assignment and the exam have been compared and particularly in similar subject areas for example requirements analysis and software architecture. Also, students' and tutors' feedback have been analyzed.

Noises in images: Image noise is random (not present in the object imaged) variation of brightness or color information in images and is usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera. Image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector. Image noise is an undesirable by-product of image capture that adds spurious and extraneous information.

Salt and pepper noise: Fat-tail distributed or "impulsive" noise is sometimes called salt and pepper noise or spike noise. An image containing salt-and-pepper noise will have dark pixels in bright regions and bright pixels in dark regions. This type of noise can be caused by analog to digital converter errors, bit errors in transmission, etc. It can be mostly eliminated by using dark frame subtraction and interpolating around dark/bright pixels.

$$f(x) = \begin{cases} f_a & \text{for } x = a \\ f_b & \text{for } x = b \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

In Eq. 1, if f_a or f_b is zero, researchers have unipolar impulse noise. If both are nonzero and almost equal, it is called salt and pepper noise. Impulsive noises can be positive and/or negative. It is often very large and can go out of the range of the image. It appears as black and white dots or saturated peaks. Salt and pepper noise is a form of noise typically seen on images. It represents itself as randomly occurring white and black pixels. An effective noise reduction method for this type of noise involves the usage of a median filter, morphological filter or a contra harmonic mean filter. Salt and pepper noise creeps into images in situations where quick transients such as faulty switching, take place.

Amplifier noise (Gaussian noise): The standard model of amplifier noise is additive, Gaussian, independent at each pixel and independent of the signal intensity, caused primarily by Johnson-Nyquist noise (thermal noise) including that which comes from the reset noise of capacitors ("kTC noise"). Amplifier noise is a major part of the "read noise" of an image sensor that is of the constant noise level in dark areas of the image. In color cameras where more amplification is used in the blue color channel than in the green or red channel, there can be more noise in the blue channel:

$$I_v(x, y) = I(x, y) + V(x, y) \tag{2}$$

Where:

I_v = The observed image with noise
I = The true signal (image)
V = The noise component

Many Additive Noise Models exist and the following are some common Additive Noise Models with their Probability Density Function (PDF).

Gaussian noise is statistical noise that has its probability density function equal to that of the normal distribution which is also known as the Gaussian distribution. In other words, the values that the noise can take on are Gaussian-distributed. A special case is white Gaussian noise in which the values at any pair of times are identically distributed and statistically independent (and hence uncorrelated). In applications, Gaussian noise is most commonly used as additive white noise to yield additive white Gaussian noise:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{3}$$

Where:

- μ = The mean and it determines the location of the peak of the density function
- σ = Standard deviation
- σ^2 = Variance of the distribution

Uniform noise: Uniform noise is not often encountered in real-world imaging systems but provides a useful comparison with Gaussian noise. The PDF of uniform distribution is given by Gonzalez and Woods (2007):

$$f(x)=\begin{cases}\dfrac{1}{\sigma^2\sqrt{3}} & \text{for } |x|\le\sigma\sqrt{3}\\ 0 & \text{else}\end{cases} \tag{4}$$

MATERIALS AND METHODS

Mutiwavelet transform: Multi-wavelets were defined using several wavelets with several scaling functions (Santini and Jain, 1999). Multi-wavelets have several advantages in comparison with scalar wavelet (Hiremath and Shivashankar, 2006). The features such as compact support, orthogonality, symmetry and high order approximation are the base features for this transformation. A scalar wavelet cannot possess all these properties at the same time. On the other hand, a multiwavelet system can simultaneously provide perfect representation while preserving length (Orthogonality), good performance at the boundaries (via linear-phase symmetry) and a high order of approximation (vanishing moments) (Wan and Kuo, 1996). Thus, multiwavelets offer the possibility of superior performance and high degree of freedom for image processing applications, compared with scalar wavelets.

These multiwavelets have the potential to offer better representative quality than the conventional scalar transforms. Finally, multiwavelets can achieve better level of performance than scalar wavelets with similar computational complexity. Wavelets are useful tools for signal processing applications such as image retrieval and denoising.

During a single level of decomposition using a scalar wavelet transform, the 2D image data is replaced by four blocks corresponding to the sub bands representing either low pass or high pass in both dimensions. These sub bands are illustrated in Fig. 1 and 2.

The multi-wavelets used here have two channels, so there will be two sets of scaling coefficients and two sets of wavelet coefficients. Since, multiple iteration over the low pass data is desired, the scaling coefficients for the two channels are stored together.

Likewise, the wavelet coefficients for the two channels are also stored together. For multi-wavelets the L and H have subscripts denoting the channel to which the data

LL	LH
HL	HH

Fig. 1: Scalar wavelets

L1L1	L1L2	L1H1	L1H2
L2L1	L2L2	L2H1	L2H2
H1L1	H1L2	H1H1	H1H2
H2L1	H2L2	H2L1	H2L2

Fig. 2: Multi-wavelet

corresponds. For example, the sub-band labeled L1H2 corresponds to data from the second channel high pass filter in the horizontal direction and the first channel low pass filter in the vertical direction. This shows how a single level of decomposition is done. In practice, there is more than one decomposition performed on the image. Successive iterations are performed on the low pass coefficients from the pervious stage to further reduce the number of low pass coefficients. Since, the low pass coefficients contain most of the original signal energy, this iteration process yields better energy compaction. After a certain number of iterations, the benefits gained in energy compaction becomes rather negligible compared to the extra computational effort. Usually, five levels of decomposition are used.

Global features representation: Compute the Standard Deviation (SD), energy and mean of the multi wavelet decomposed image. Standard deviation is:

$$\sigma_k=\sqrt{\frac{1}{m\times n}\sum_{i=1}^{m}\sum_{j=1}^{n}(w_k(i,j)-\mu_k} \tag{6}$$

Where:

- W_k = Co-efficients of kth multiwavelet decomposed sub-band
- μ_k = Mean value of kth sub-band
- $M\times N$ = Size of the multiwavelet decomposed sub-band

Energy is:

$$E_k = \frac{1}{m \times n}\sum_{i=1}^{m}\sum_{j=1}^{n}\left|w_{ij}\right| \tag{7}$$

HSV colour space and quantization: HSV are the two most common cylindrical-coordinate representations of points in an RGB Color Model. The two representations rearrange the geometry of RGB in an attempt to be more intuitive and perceptually relevant than the cartesian (cube) representation. The hue (H) of a color refers to which pure color it resembles. All tints, tones and shades of red have the same hue. Hues are described by a number that specifies the position of the corresponding pure color on the color wheel, as a fraction between 0 and 1. Value 0 refers to red; 1/6 is yellow; 1/3 is green and so forth around the color wheel. The Saturation (S) of a color describes how white the color is. A pure red is fully saturated with a saturation of 1; tints of red have saturations <1 and white has a saturation of 0. The Value (V) of a color, also called its lightness, describes how dark the color is. A value of 0 is black with increasing lightness moving away from black. The outer edge of the top of the cone is the color wheel with all the pure colors. The H parameter describes the angle around the wheel. The S (saturation) is zero for any color on the axis of the cone; the center of the top circle is white. An increase in the value of S corresponds to a movement away from the axis. The V (Value or lightness) is zero for black. An increase in the value of V corresponds to a movement away from black and toward the top of the cone (Fig. 3).

MSD definition and local feature representation: Human visual system is sensitive to orientation and color. Orientation is a powerful visual cue about the subject depicted in an image. Strong orientation usually indicates a definite pattern, however, many natural scenes do not show strong orientation and have no clear structure or specific pattern. Although, the natural images show various contents, they may have some common fundamental elements. The different combination and spatial dis-tribution of those basic elements result in the various micro-structures or patterns in the natural images. In this study, micro-structures are defined as the collection of certain underlying colors. The underlying colors are those colors which have similar or the same edge orientation in uniform color space. The highlight of underlying colors is that they can combine color, texture and shape cues as a whole. Julesz's Texton theory mainly focuses on analyzing regular textures while the micro-structures can be considered as the extension of Julesz's textons or the color version of textons. Since, micro-structures involve color, texture and shape information, they can better present image features for image retrieval.

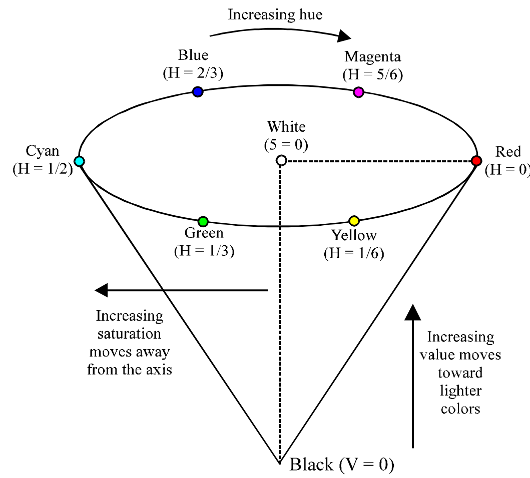

Fig. 3: HCV colour space

In order to find the micro-structures (Liu *et al*., 2011) which have similar attributes such as edge orientation and color distribution, researchers partition the image into many small blocks which can be a grid of size 2×2, 3×3, 5×5, 7×7 and so on. For the convenience of expression, the 3×3 block is used in the following development of micro-structure analysis. The edge orientation image y(x, y) is used to define micro-structures because an edge orientation is insensitive to color and illumination variation and it is independent of translation, scaling and small rotation. Note that since, researchers quantize the orientation into six levels, the values of the pixels in y (x, y) can only vary from 0-5. Researchers move the 3×3 block from left to right and top to bottom throughout the image to detect micro-structures.

Values of a micro-structure image f(x,y) is denoted as $f(x, y) = w, w \in \{0, 1, ..., L-1\}$. SIn each 3×3 block of f(x, y), denoted by $P_0 = (x_0, y_0)$ the center position of it and let f $(P_0) = W_0$ denoted by $P_i = (X_i, Y_i)$ eight nearest neighbours to P_0 and let f $(P_i) = w_i$, i = 1, 2, ..., 8. Denoted by N the co-occurring number of values w_0 and w_i and by N the occuring number of w_0. Moving 3×3 block from left to right and top to bottom throughout the micro-structure image, researchers use the following Eq. 8 to describe the micro-structure features:

$$H(w_0) = \begin{cases} \dfrac{N\{f(p_0) = w\_(0) \wedge f(p_i) = w_i \,|\, |p_i - p_0| = 1\}}{8N\{f(p_0) = w_0\}} \\ \text{where} \qquad w_0 = w_i, i \in \{1, 2, \ldots, 8\} \end{cases} \tag{8}$$

The dimensionality of (H_{w0}) is 72 for color images. It can express how the spatial correlation of neighbouring underlying colours distributes in the micro-structures image.

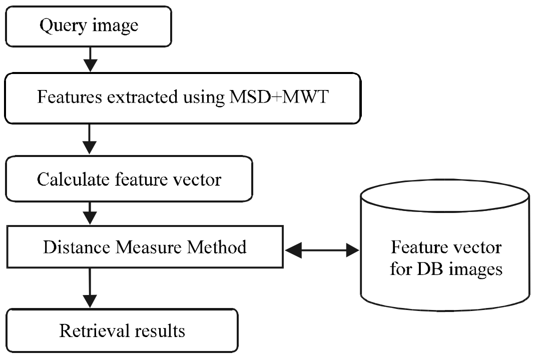

Fig. 4: Proposed method

Combined Global-Local Specialized Feature Proposed algorithm:

- Step 1: Query image convert to RGB to HSV
- Step 2: Calculate MSD features using Eq. 8
- Step 3: Decompose query image in the multi-wavelet domain
- Step 4: Calculate SD, energy and mean using Eq. 6 and 7
- Step 5: Apply query image and calculate the feature vector as given in steps 2 and 4
- Step 6: Calculate the similarity measure using SAD and L1 distance metric
- Step 7: Retrieve all relevant images to query image based on minimum SAD or L1 distance metric (Fig. 4)

RESULTS AND DISCUSSION

Data set: So, far there are no standard test data set and performance evaluation model for CBIR Systems (Liu *et al*., 2007). Most of the researchers use corel image data set to test image retrieval performance while some researchers use self-collected images or Brodatz and Outex texture data sets in experiments. However, corel dataset has become a de-facto standard in demonstrating the performance of CBIR Systems.

Corel image database contains a large amount of images of various contents ranging from animals and outdoor sports to natural scenarios. Image size is 125×125 in the jpeg format. The database contains 20 categories with total 500 images. Sum of Absolute Differences (SAD) Method is used.

Similarity measure

Sum of Absolute Differences (SAD): In CBIR, eight similarity measures have been proposed (Wan and Kuo, 1996). In this research, Sum of Absolute Differences (SAD) Method is used:

$$SAD(f_p, f_t) = \sum_{i=0}^{n} \left| f_{q[i]} - f_{t[i]} \right| \quad (9)$$

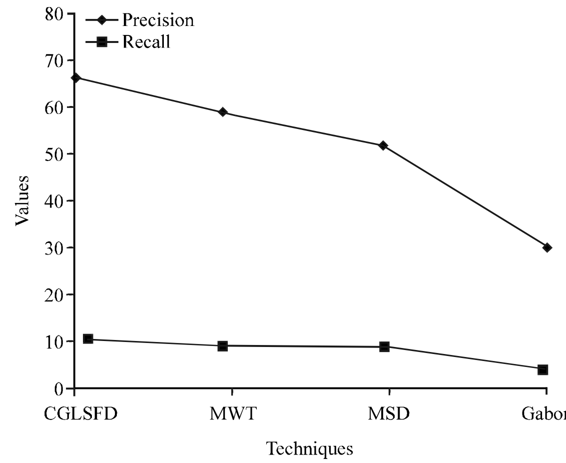

Fig. 5: Precision and recall values impulsive noise

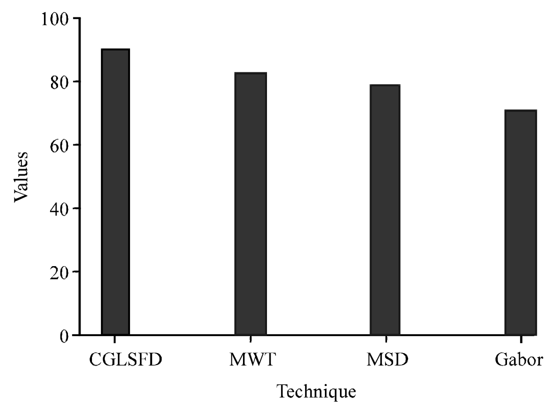

Fig. 6: Performance rate comparsion of query with impulisve noise

where, $f_q[1]$, $f_q[2]$, ..., f_q [n] represents the query feature vector, f_t [1], f_t [2], ..., f_t [n] denotes the database feature vectors and n is the number of features in each vector.

L1 distance: An = M-dimensional feature vector $T = [T_1, T_2, ..., T_M]$ is extracted and stored in the database. Let $Q = [Q_1, Q_2, ..., Q_M]$ be the feature vector of a query image, the L_1distance between them is simply calculated is:

$$D(T, Q) = \sum_{i=1}^{M} \left| T_i - Q_i \right| \quad (10)$$

The L1 distance is simple to calculate which needs no square or square root operations. It can save much computational cost and is very suitable for large scale image datasets. For the proposed MSD, M = 72 for color images. The class label of the template image which yields the smallest distance is assigned to thequery image (Fig. 5 and 6 and Table 1-3).

Table 1: Comparison of precision and recall value for CGLSFD, MWT, MSD and Gabor

	Techniques			
Performance	CGLSFD	MWT	MSD	Gabor
Precision	67.33	59.20	55.92	38.26
Recall	8.56	7.12	6.78	4.38

Table 2: Comparison of SAD and L1 distance metric value for CGLSFD, MWT, MSD and Gabor

	Techniques			
Distance metric	CGLSFD	MWT	MSD	Gabor
SAD	82.23	75.37	68.40	52.83
L1	63.78	58.37	55.92	51.27

Table 3: Performance rate comparison of noisy query in methods CGLSFD, MWT, MSD and Gabor

	Techniques			
Noise details	CGLSFD	MWT	MSD	Gabor
Impulsive noise	90.56	82.23	78.45	70.21
Additive	53.89	41.71	31.58	26.33

CONCLUSION

In this study, researchers introduced a novel CBIR approach via selected specialized features in multi-wavelet decomposition and MSD of CGLSFD in images, followed by CGLSFD feature extraction and similarity match under SAD and L1. The precision value goes on increasing and recall value goes on decreasing as decomposition level increases. Also, the proposed method perform well in various noisy query environment, CGLSFD performance is good in compared with some traditional existing methods Gabor, MSD and MWT.

REFERENCES

Alnihoud, J., 2012. Content-based image retrieval system based on self organizing map, fuzzy color histogram and Subtractive fuzzy clustering. Int. Arab J. Inform. Technol., 9: 452-458.

Daubechies, I., 1990. The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. Inform. Theo., 36: 961-1005.

Desimone, R., 1998. Visual attention mediated by biased competition in extrastriate visual cortex. Philos. Trans. R. Soc. London B: Biol. Sci., 353: 1245-1255.

Do, M.N. and M. Vetterli, 2002. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. IEEE Trans. Image Process, 11: 146-158.

Gonzalez, R.C. and R.E. Woods, 2007. Digital Image Processing 3rd Edn., Prentice Hall, New York, ISBN-13: 9780131687288.

Hiremath, P.S. and S. Shivashankar, 2006. Wavelet based features for texture classification. J. Graphics Vision Image Process., 6: 55-58.

Liu, G.H., Z.Y. Li and L. Zhang and Y. Xu, 2011. Image retrieval based on micro-structure descriptor. Pattern Recog., 44: 2123-2133.

Liu, Y., D. Zhang, G. Lu and W.Y. Ma, 2007. A survey of content-based image retrieval with high-level semantics. Pattern Recogn., 40: 262-282.

Mallat, S.G., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell., 11: 674-693.

Santini, S. and R. Jain, 1999. Similarity measures. IEEE Trans. Pattern Anal. Mach. Intell., 21: 871-883.

Sonka, M., V. Hlavac and R. Boyle, 2008. Digital Image Processing and Computer Vision. Cengage Learing, USA.

Wan, X. and C.C.J. Kuo, 1996. Color distribution analysis and quantization for image retrieval. Proceedings of SPIE Conference on Storage and Retrieval for Still Image and Video Databases IV, Volume 2670, January 28-February 2, 1996, San Jose, CA., USA., pp: 8-16.

Zhang, D. and G. Lu, 2003. Evaluation of similarity measurement for image retrieval. Proceedings of the IEEE International Conference Neural Networks and Signal Processing, December 14-17, 2003, Nanjing, China, pp: 928-931.